MPEG-4 Demystified

Session 702

MPEG-4 Demystified

Aimee Nugent
Sr. Product Manager, QuickTime
nugent@apple.com

MPEG-4 Demystified

Rob Koenen
President, MPEG-4 Industry Forum
Chairman, MPEG Requirements Group
Vice President, InterTrust Technologies
rob.koenen@m4if.org

Overview

- What is MPEG-4?
- How does MPEG-4 work?
- Recent developments in MPEG-4
- Why use MPEG-4?
- Deployment of MPEG-4
- About M4IF

Overview

- What is MPEG-4?
- How does MPEG-4 work?
- Recent developments in MPEG-4
- Why use MPEG-4?
- Deployment of MPEG-4
- About M4IF

What Is MPEG-4?

- The Media Standard
  - Set by ISO/IEC
- An interoperable cross-platform ecosystem
- An architecture and coding methods for representing rich multimedia content
  - Beyond video and audio, supporting synthetic content, raster and vector graphics
- Designed for **all** digital multimedia platforms

Where Does It Come from?

From Moving Picture Experts Group, who also did:

- MPEG-1: Standard for Video CD, CD-I, web, digital photo cameras
  - MPEG-1 Layer III audio codec (MP3)
- MPEG-2: the DVD and digital TV standard

And also:

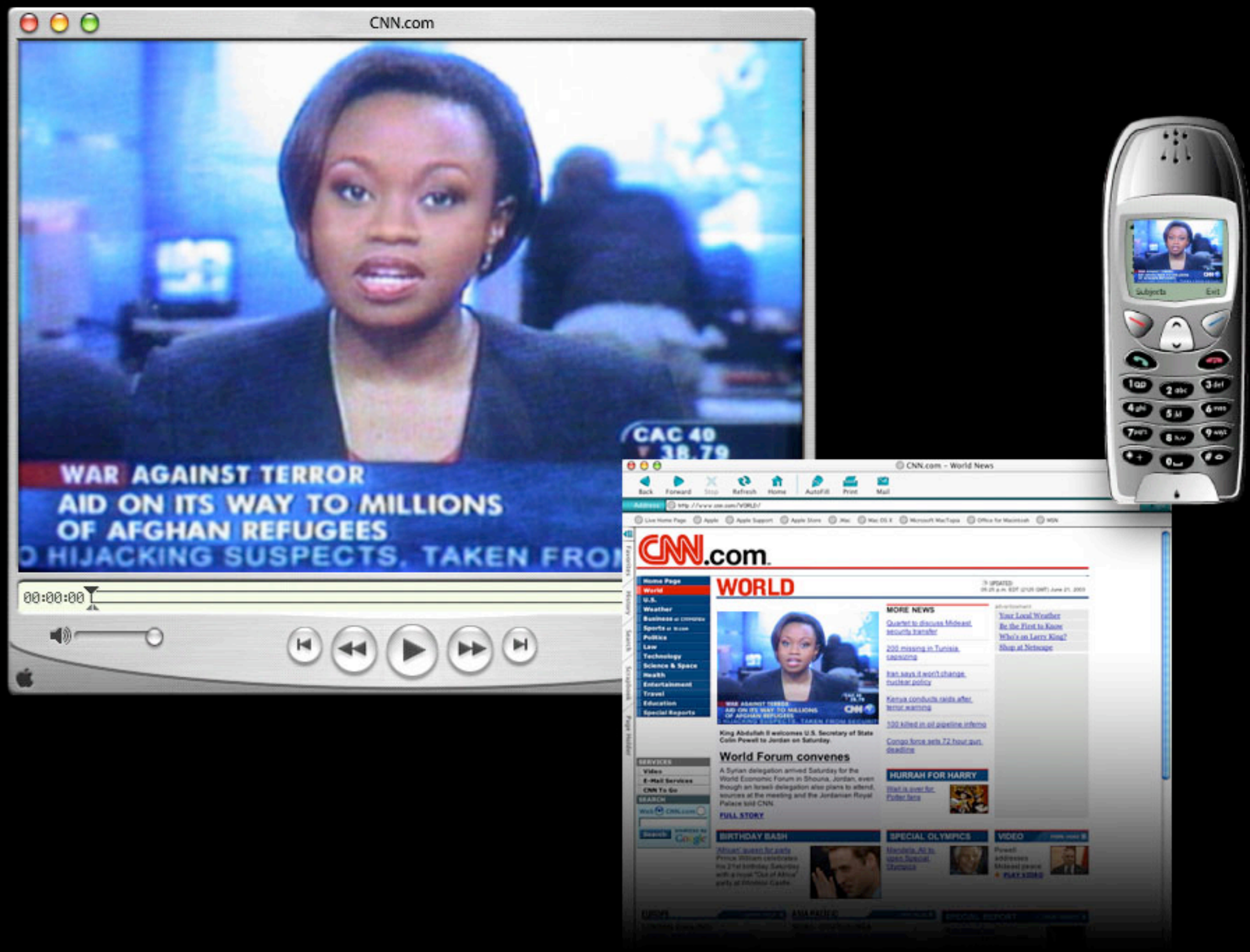
- MPEG-7: standard for description of content
- MPEG-21: framework for interoperable use and exchange of digital media

MPEG-4 Vision

- No **convergence**, but **proliferation** of multimedia over different networks, terminals
- **Common** multimedia technology necessary that supports:
  - Broadcast
  - Communication
  - Retrieval
    - Online
    - Packaged media

One Single Technology for…

Write Once—Play Everywhere



Applications of MPEG-4

- Mobile devices
  - Low datarates, errors, scalability
- Broadcast
  - New, on-demand services
  - HD broadcast
  - “MPEG-4 over MPEG-2”

- Streaming services
  - Scalability, low to medium datarates, interactivity
- Packaged Media (DVD)
  - Interactivity
  - HD through higher compression rates

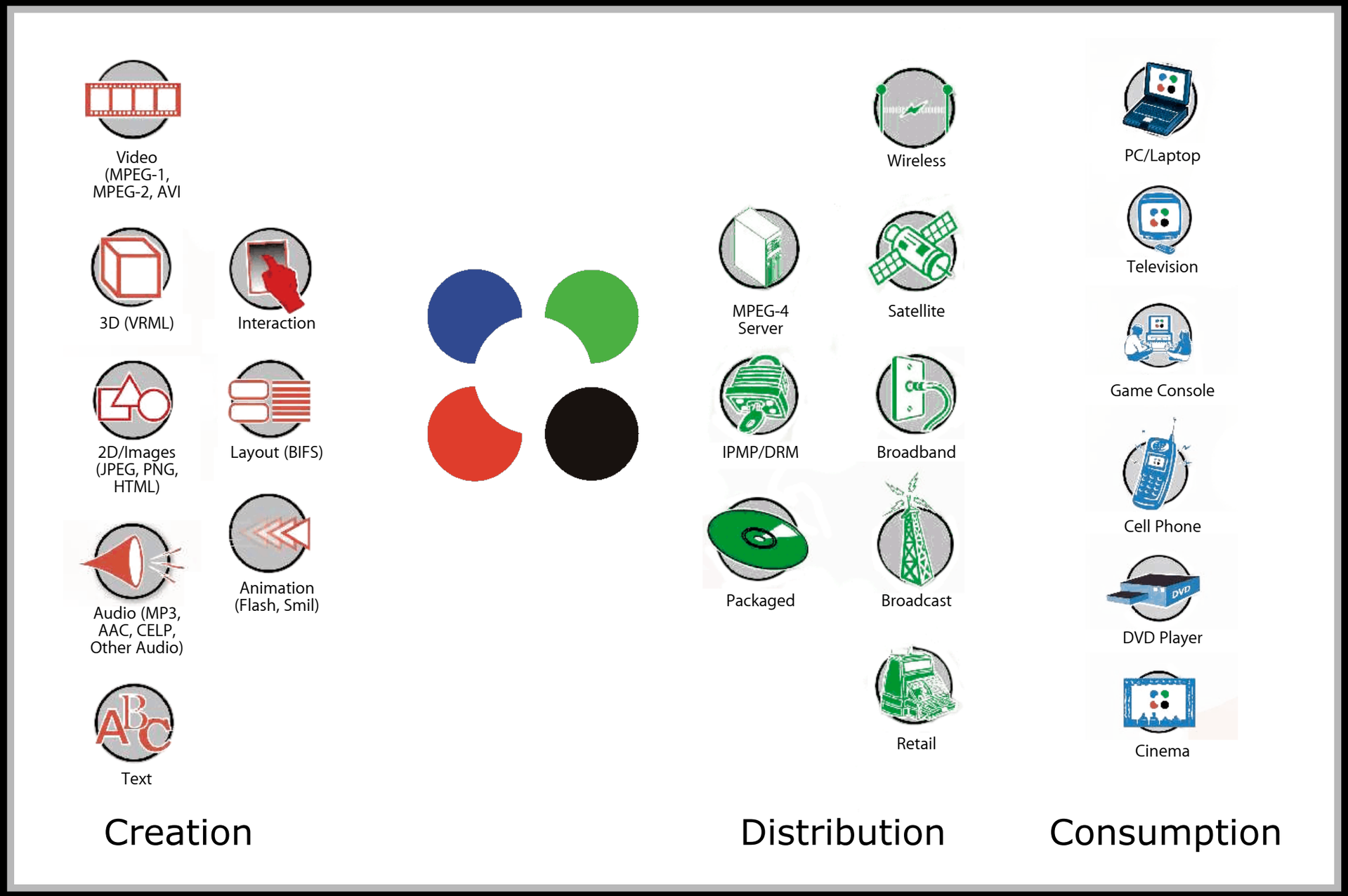
Overview

- What is MPEG-4?
- How does MPEG-4 work?
- Recent developments in MPEG-4
- Why use MPEG-4?
- Deployment of MPEG-4
- About M4IF

The Nature of MPEG-4

- Object-based multimedia content representation standard
- Revolutionary systems layer
- State-of-the art coders, responsibly upgraded
- Profiles and levels restrict complexity and guarantee interoperability

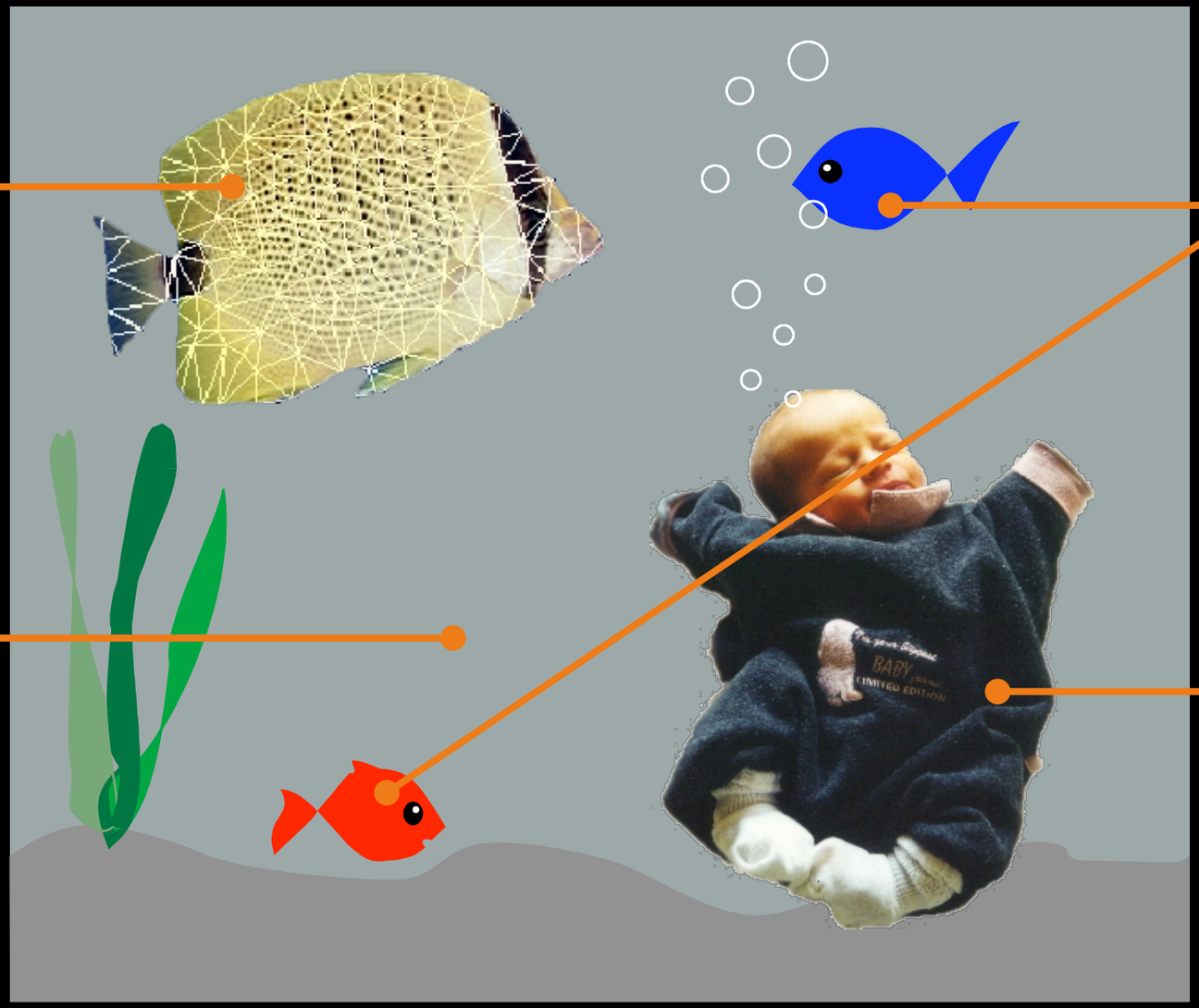
The MPEG-4 Ecosystem



An MPEG-4 Scene

Animated 2D Photo / Mesh

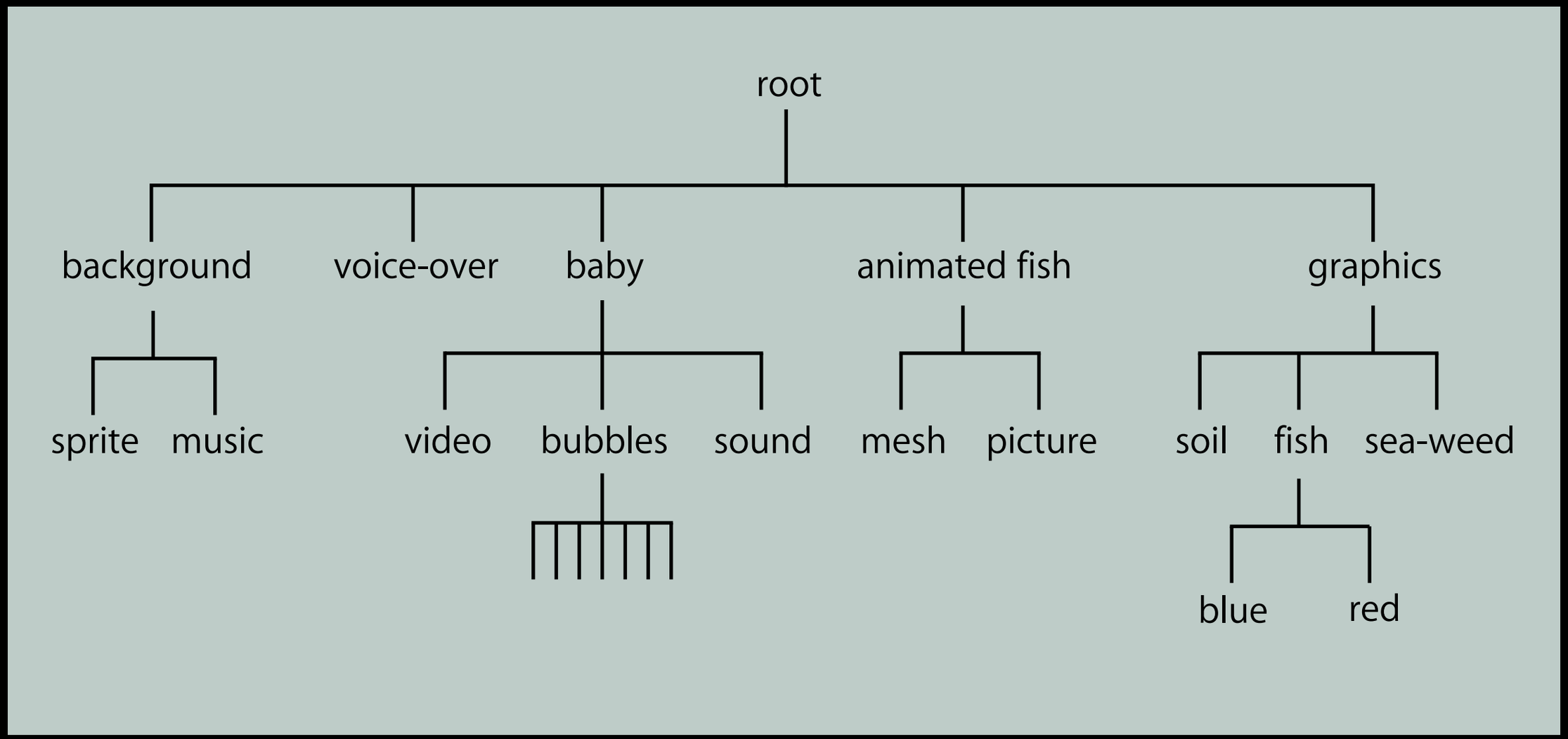
Graphic Objects



Background with Music

'Natural' Audiovisual Object

The Scene Tree



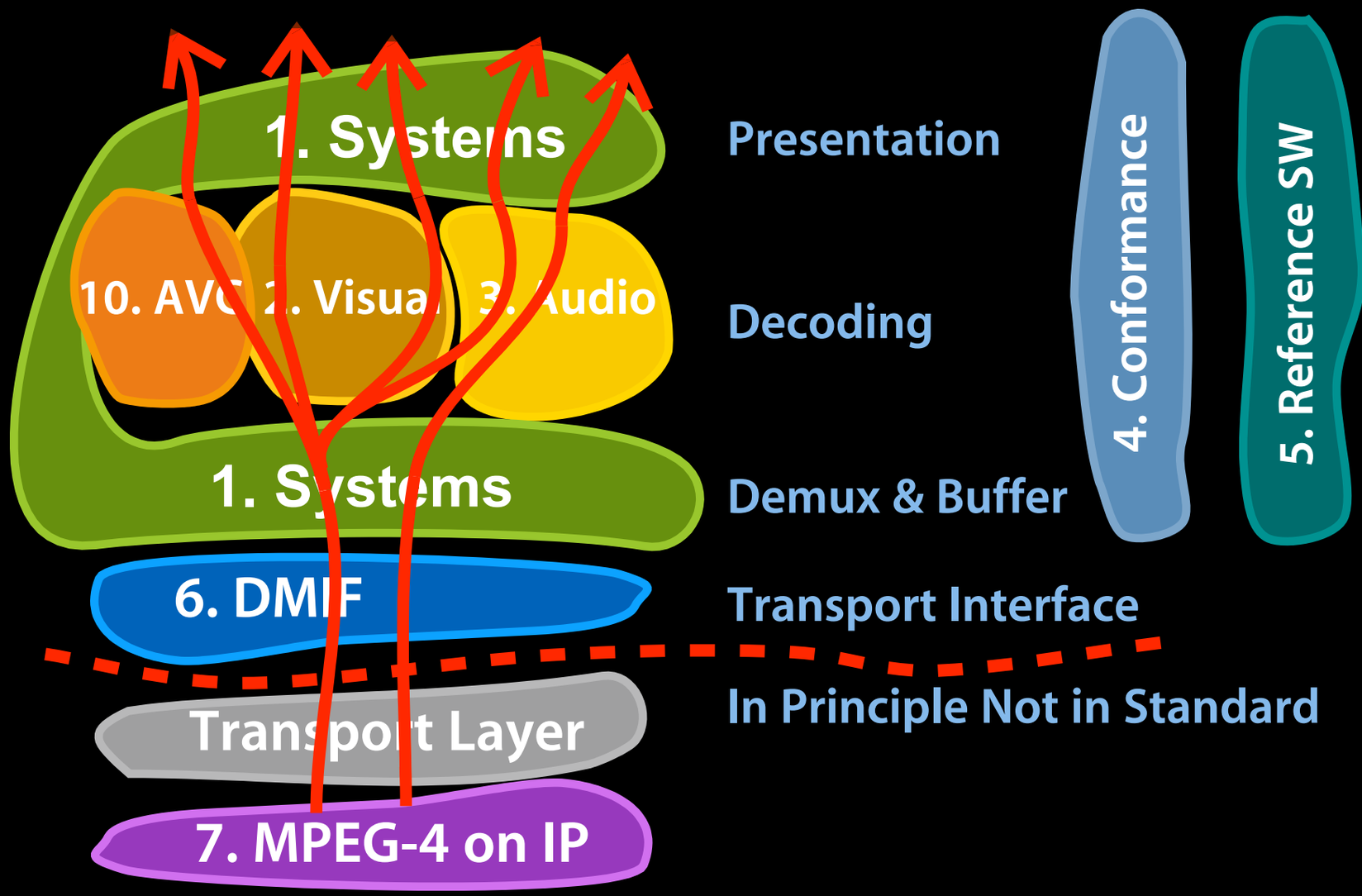
Recapping: MPEG-4 Is About the Coding of Audiovisual Objects

- Audiovisual **Scene** is with "objects"
- These **Objects** can be of different nature
- **Compositor** puts objects in scene
- Efficient, real-time binary **scene description language** (BIFS)
- **Coding** scheme is optimal for object type
- Principle is fully **independent** of bitrate

MPEG-4 Supports a Variety of “Objects”

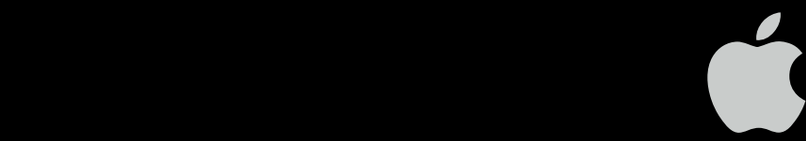
- Visual
  - Video
  - Animated face and body
  - 2D and 3D animated meshes
  - Text and graphics
- Audio
  - General audio—mono—5.1 channels
  - Speech
  - Synthetic sounds (”Structured Audio”)
  - Text-to-speech
  - “Environmental spatialization”

The Parts of the MPEG-4 Standard



Visual Media Object Types in MPEG-4

- Video from 10–1,000,000 Kbit/s
  - Multiple rectangular or arbitrary shape objects in the scene
  - Scalable—Interlaced and Progressive
- “Sprites” (e.g. backgrounds): send once, then warp
- Computer-generated visual information
  - Synchronized graphics and animated text
  - Face and body animation
  - Animated meshes with moving texture



Audio Media Objects in MPEG-4

- Audio from 4 kbit/s to 64 kbit/s/channel
  - Arbitrary number of objects in the scene
  - MPEG-4 AAC, TWINVQ
  - > 4kbit/s: Harmonic and Individual Lines plus Noise' (HILN)
- Voice from 2 kbit/s to 24 kbit/s
  - 2-4 kbit/s: HVXC (Harmonic Vector eXcitation Coding)
  - 4-24 kbit/s: CELP:
- Large step + fine grain scalability
  - 1 kbit/s steps through bit-sliced arithmetic coding

MPEG-4's Synthetic Audio Objects

- Structured audio
  - SAOL (SA Orchestra Language)
  - SASL (SA Score Language)
  - Great "music" at very low bitrates
- Coded form of MIDI
- Wavetable synthesis for simple decoders
  - Including effects
- Text-to-speech (interface!)
  - To complement face/body animation

MPEG-4 Systems

- Binary Scene Description
  - VRML concepts + streaming + real time + efficiency
  - Support for content in scene from different sources
  - Allows interaction (local/remote)
  - 2D and 3D
  - Dynamic scene updates and scene animation
- XMT: eXtensible MPEG-4 textual format
  - Textual format for BIFS
  - Includes some SMIL harmonization

MPEG-4 Systems

- Predictable behavior of a decoder and decoded content:
  - Tight synchronization of A, V, synthetic, graphic elements
  - Buffer management
- Flextime
  - Spring-like timing model (a-la SMIL)
- IP management and protection (IPMP) Interface
  - Standard interface to proprietary DRM Systems
  - MPEG-21 to bring more interoperability in DRM
- File format MP4 (based on QuickTime)

MPEG-4 Systems (Cont.)

- MPEG-J(ava)
  - API for complicated content behavior,
  - API's to network/terminal/UI resources
  - "Application Engine"
- Audio rendering
  - Specify downmix from arbitrary number of channels
  - Environmental spatialization
    - Modeling of environment for spatial sound reproduction
      - Physical and perceptual model

Profiles and Levels

- Conformance points are “Profiles@Levels”
  - A bit like in MPEG-2
- Profiles determine tool set
  - For example, B frames, ¼ pel Motion Compensation
- Levels limit complexity
  - For example, MacroBlocks/sec, Complexity Units in Audio
- Profiles are convergence point for Industry Standards built on MPEG-4
  - Vehicle for market decisions and uptake
  - For instance, Internet Streaming Media Alliance

Profile Dimensions:

Media Profiles

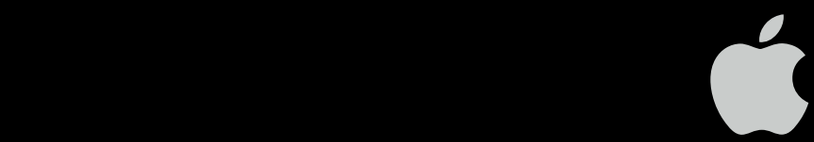
- Scene Description (Scene Graph)
  - Tools to describe and manipulate scene
- MPEG-J (Main and Personal)
- Object Descriptor (Synch and Buffers)
- MPEG does not prescribe how to combine these
  - That's what other industry fora do

Overview

- What is MPEG-4?
- How does MPEG-4 work?
- Recent developments in MPEG-4
- Why use MPEG-4?
- Deployment of MPEG-4
- About M4IF

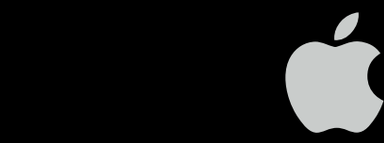
MPEG-4 Advanced Video Coding

- Also standardized as ITU-T H.264
  - Part 10 of MPEG-4 standard
  - Also known as the “JVT codec”
- Built by world’s top video coding experts from ITU and MPEG in Joint Video Team (JVT)
- Best quality/complexity trade-off
  - Better than WM9, according to industry leader LSI Logic
  - Improvements will continue
- Significantly further reduces video bitrates
- Fits into the MPEG-4 framework
  - Will coexist with MPEG-4 part 2 in market

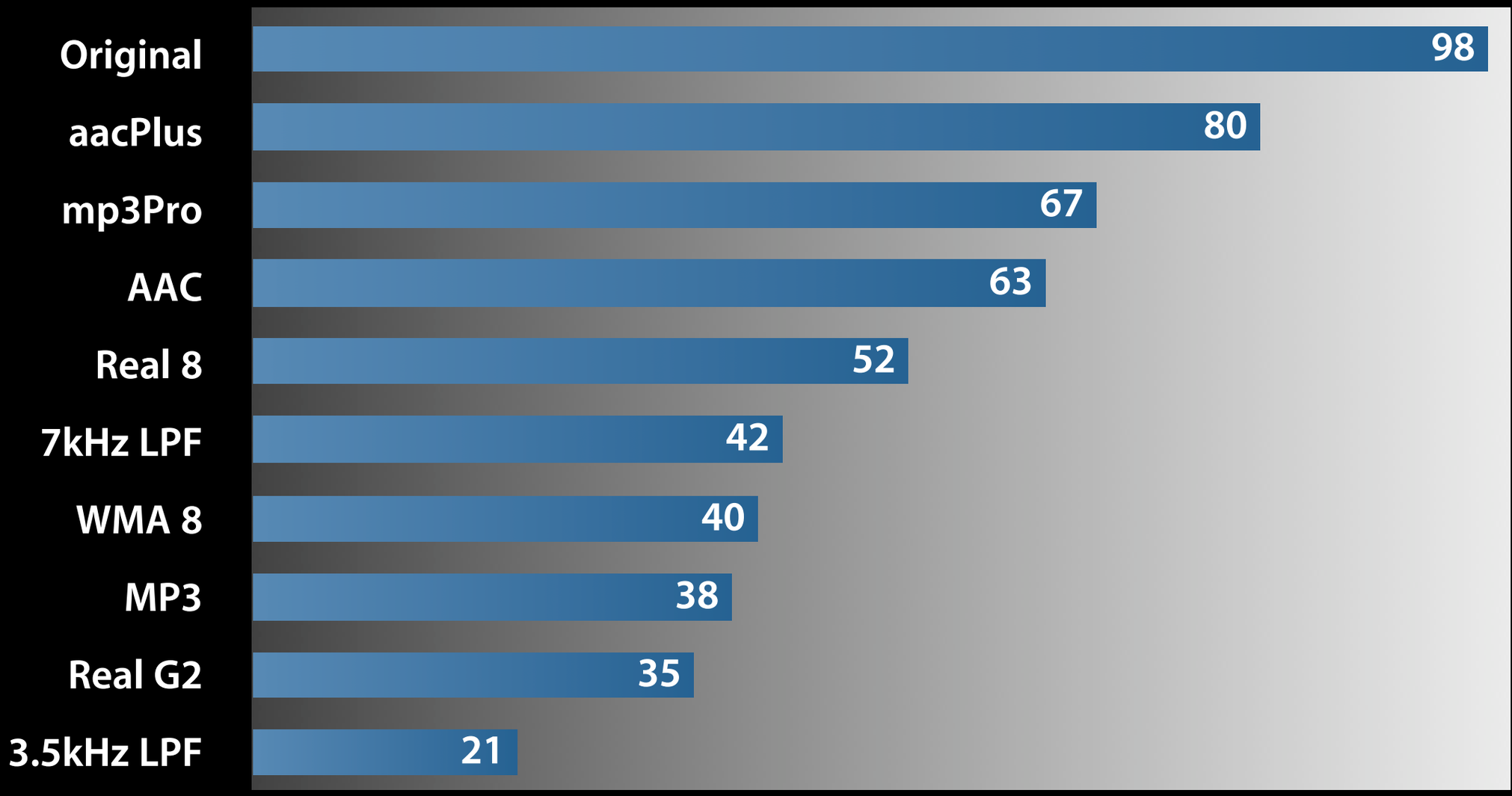


High-Efficiency Advanced Audio Coding

- CD quality @ 48kbps; high quality @ 32kbps
  - Achieved through “Spectral Bandwidth Replication”
  - Predict upper half of spectrum from lower half
  - Forward and backward compatible with “normal” AAC
- Amendment (addition) to MPEG-4 Audio
- Already operational in XM radio, chosen for Digital Radio Mondiale (“DRM”)
- Tested as best codec by EBU (European Broadcasting Union) over Windows, Real



EBU Listening Test Results for 48Kbps Stereo



LPF = Low Pass Filter.
aacPlus is brand name for High Efficiency AAC.
mp3PRO is brand name for MPEG-1 Layer III with Spectral Band Replication
EBU = European Broadcasting Union

Other Developments

- Animation Framework eXtension (AFX)
  - Unified framework for interactive animation and graphics
- Work on 3D video coding is in requirements phase
- Work on lossless audio coding is entering development phase

Overview

- What is MPEG-4?
- How does MPEG-4 work?
- Recent developments in MPEG-4
- Why use MPEG-4?
- Deployment of MPEG-4
- About M4IF

Standards Make Sense—They…

…fuel **innovation** (GSM, WLAN/802.11, …)

…**survive** the test of time (PAL, NTSC, MP3, …)

…prevent expensive **format churn** (VHS, CD, …)

…help markets **mature** (CD, DVD, MPEG-2, …)

…provide an **interoperable ecosystem** of tools and equipment—different providers make what they excel in

- No **vendor lock-in**—avoid channel conflicts
- Market-based **price control** mechanisms
- **Multi-vendor** sourcing of equipment
- No monopolies—no single vendor **controls** the format

MPEG-4 Has Clear Benefits

- Code once, use across platforms and players
- Users pick their favorite players, devices
- Content producers pick their favorite tools
- Providers only need to stream in one format
- Competition drives quality up

(R)evolution

- A **revolution** in functionality
  - Designed from the start for interactivity
  - State-of-the-art compression; responsible upgrades
  - Object-based = intrinsic flexibility
  - Profiles, levels limit complexity, guarantee interoperability
  - Offers synthetic content
  - Local rendering of 2D and 3D graphics and audio
- An **evolution** in infrastructure
  - Allows all types of “casting”: unicast, multicast, broadcast…
  - On all network types, no need for new networks or transports
  - Leverage existing MPEG-2 and other infrastructures

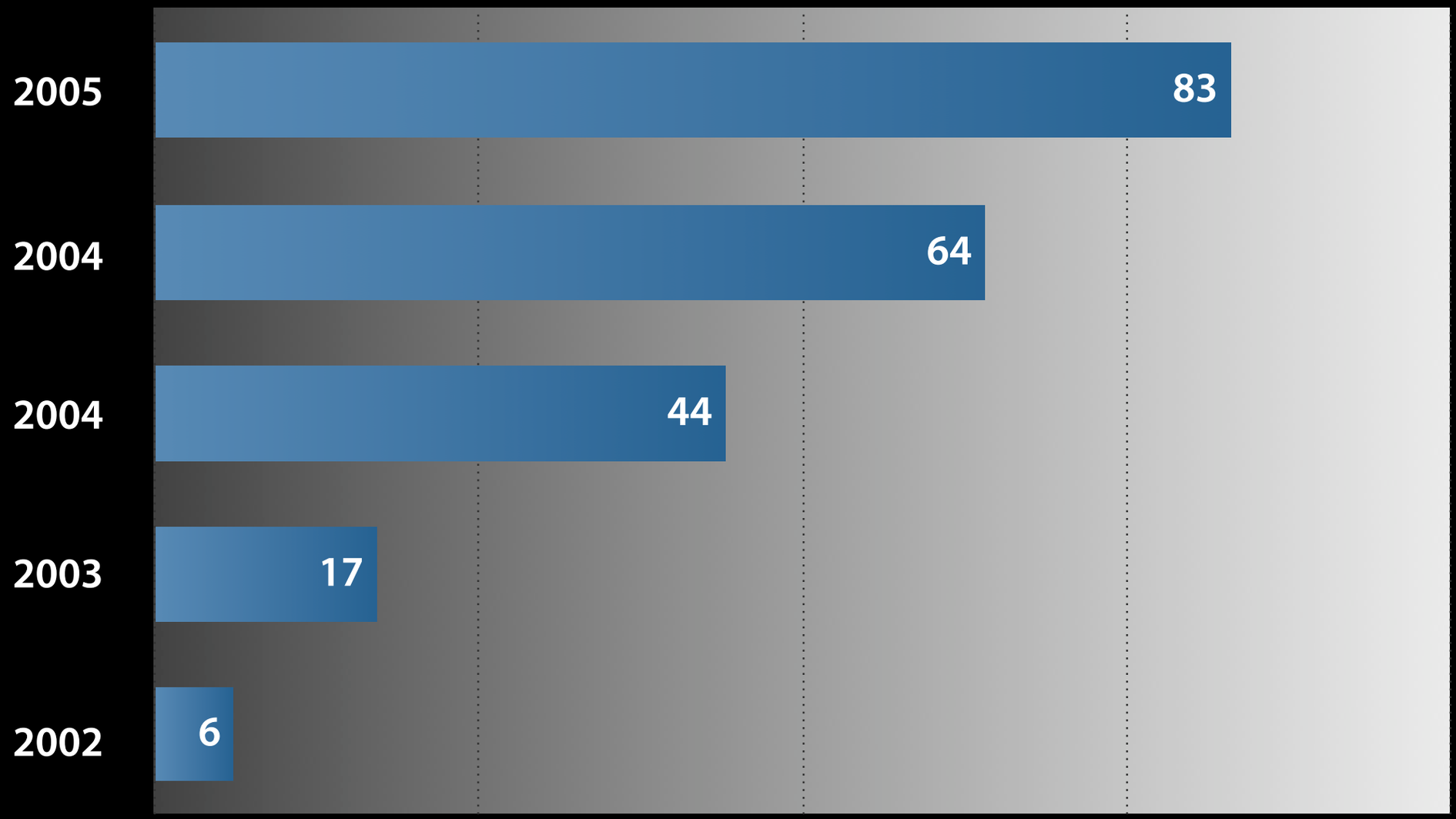
Business Benefits

- MPEG-4 **saves** money
  - More efficient use of bandwidth
  - Repurpose existing content
  - No need to duplicate work when adding value to assets
  - Integrate into existing MPEG delivery environments
  - Used on all types of network, including IP
- MPEG-4 **makes** money
  - Use existing content and infrastructure in new ways
  - Add new dimensions to content, including interactivity
  - Low-risk/high-return path for new entrants

Risks of Proprietary Technology

- Third-party business and pricing models
- Proprietary, confidential third-party technology roadmaps
  - May change without warning
- Potential channel conflicts
  - Supplier also competitor
  - License includes more “bundled features" than required
- Single sourcing problems

Worldwide MPEG-4 Forecast

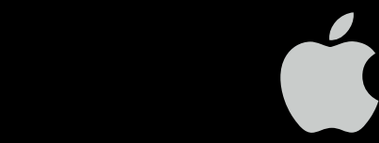


Shipments in millions of units*

*Includes standalone MPEG-4 chips and cores embedded in processors
Source: ISUPPLI CORP.

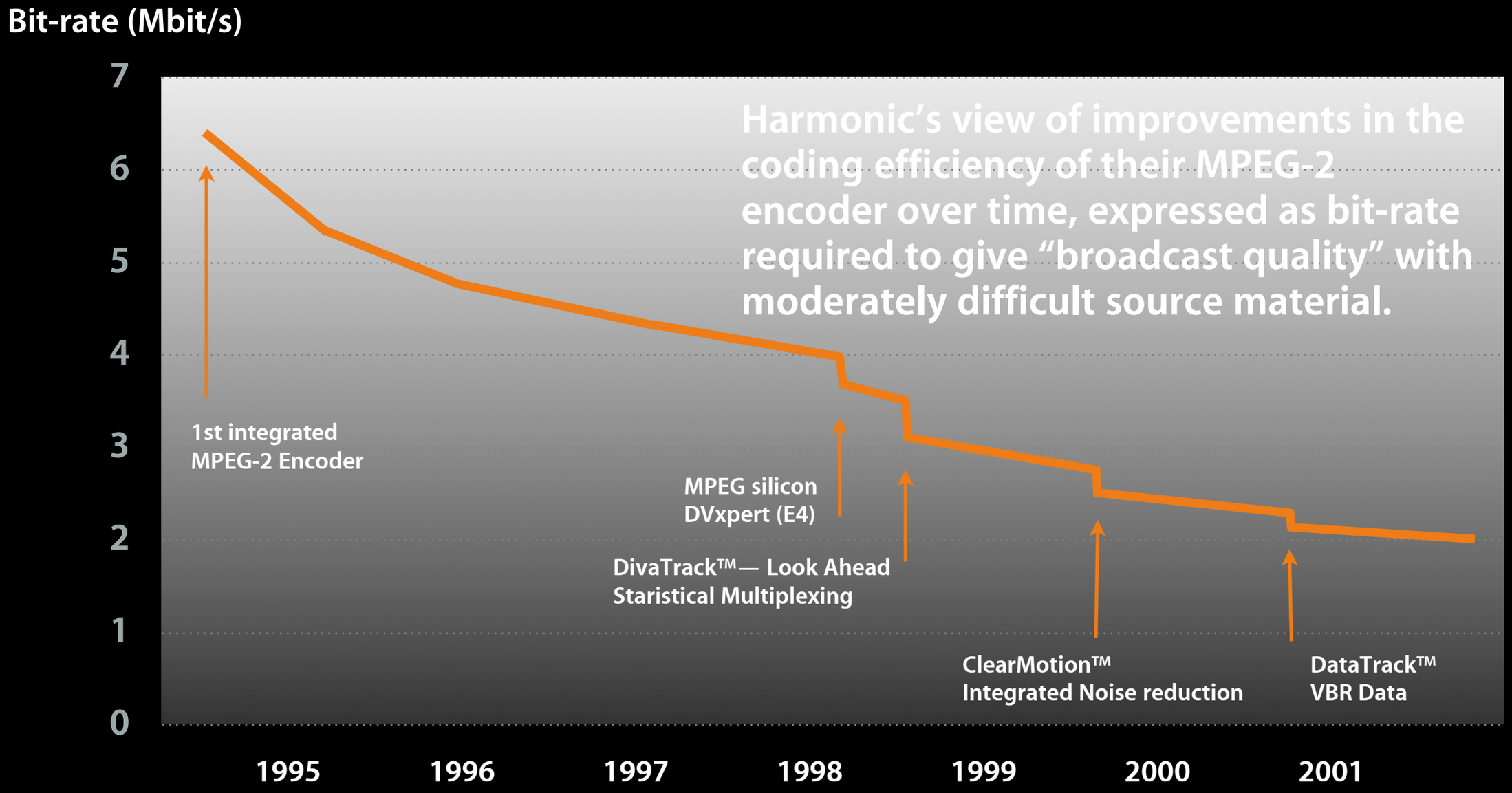
Improvement Through Competition

- Proprietary vendors underrate MPEG-4 quality:
  - Comparisons with early, non-optimized codecs
  - Industry-standard video quality does not exist!
- MPEG standardizes the minimum: only decoders
  - Competition on encoding and post-processing
- MPEG-2 bit-rates have reduced by over 50%
  - After the standard was frozen in 1996
  - Without upgrading decoders
- Competition between the “experts” will:
  - Drive MPEG-4 quality up after deployment of decoders
  - Drive functionality of tools
  - Make AVC beat all proprietary codecs, including WM9



MPEG-2 Coding Efficiency

Harmonic, Inc.'s view

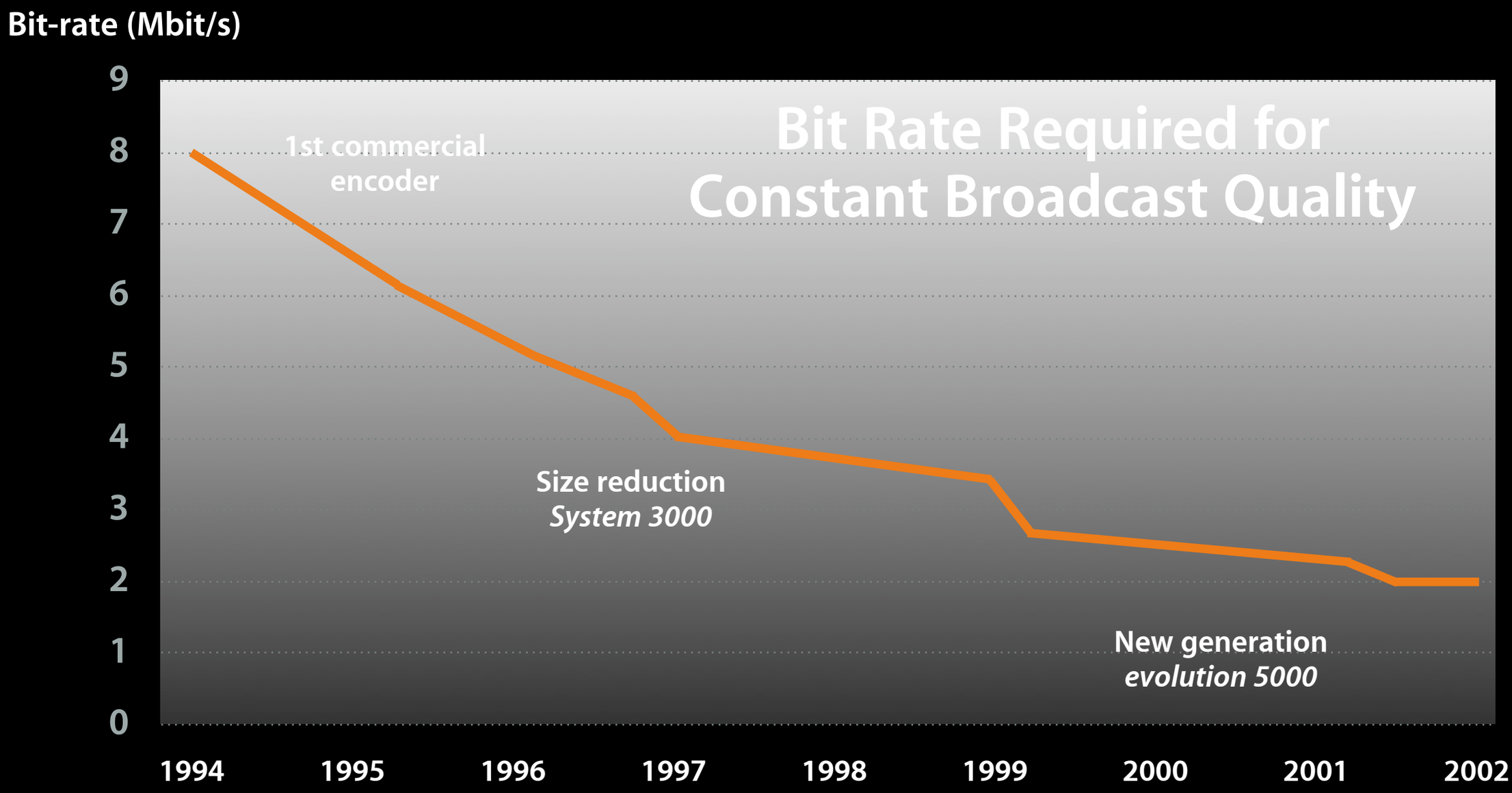


Harmonic's view of improvements in the coding efficiency of their MPEG-2 encoder over time, expressed as bit-rate required to give "broadcast quality" with moderately difficult source material.

Reproduced with the kind permission of Harmonic Inc.

MPEG-2 Coding Efficiency

TandbergTV's view



Reproduced with the kind permission of TANDBERG Television

Overview

- What is MPEG-4?
- How does MPEG-4 Work?
- Recent developments in MPEG-4
- Why use MPEG-4?
- Deployment of MPEG-4
- About M4IF

Widely Adopted and Supported

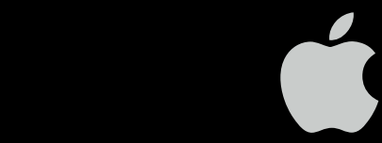
- Industry-wide PC media player support
  - QuickTime 6
  - RealOne Player/Helix (Envivio plug-in)
  - DivX: Millions of downloads weekly
  - Windows Media Player: several MPEG-4 plug-ins
- MPEG-4 is part of the standard for 3G and 2.5G mobile phones
  - 3GPP, 3GPP2 adopted MPEG-4 Profiles for video to mobile
  - AAC is the optional audio codec in 3GPP
  - File format is adaptation of MP4
  - QuickTime 6.3 supports 3GPP

Widely Adopted and Supported

- Recommended by Internet Streaming Media Alliance (ISMA)
  - Vertical spec using horizontal MPEG-4 standard
  - Founded by Sun, Kasenna, IBM, Apple, Cisco, Philips
- MPEG-4 AVC to be adopted by DVB for DVB over IP
- Becoming de facto standard for security, surveillance
- Supported in emerging Home Media Centers

Recent Announcements and Releases

- MPEG-4 HD-based video camera (Samsung)
- Solid state video cameras (Panasonic, Divio)
- DVD players that read MP4 (Toshiba, Kiss)
- HDTV Chips (Sigma Designs)
- PDAs (Sony's Clie, Sharp's Zaurus)
- Video jukeboxes (e.Digital, Archos)
- Home entertainment gateways
  (e. Digital, Equator, Sigma, …)
- Mobile phones that decode and stream MP4
  (Nokia, Motorola)



Overview

- What is MPEG-4?
- How does MPEG-4 Work?
- Recent developments in MPEG-4
- Why use MPEG-4?
- Deployment of MPEG-4
- About M4IF

About the MPEG-4 Industry Forum

History

- First Meetings 1999
  - What do we need to do to get MPEG-4 adopted?
- Officially established June 2000
  - Not-for-profit organization in Switzerland
- Now approximately 100 members
  - Worldwide
  - Across Industries—IT, Telecoms, CE, Academia, as well as content and service providers

Membership Includes

AMD
Agfa Monotype
AOL Time Warner
Apple Computer
Canon
Cisco
DivXNetworks
Dolby
Envivio
Fraunhofer
Fujitsu
Fuji-Xerox
PAL
Harmonic
Hitachi
IBM
Intel
InterTrust
iVAST
Matsushita
Microsoft
Mitsubishi
Motorola/GI
MPEG LA
NEC
NTT
Nokia
Oki
PacketVideo
Philips
RealNetworks
Scientific-Atlanta
Siemens
Sony
STMicroelectronics
Sun
TandbergTV
TDK
Tektronix
Thomson
Toshiba
Xilinx

...and many others (~100 members)

Goal

- Getting MPEG-4 widely adopted
- Doing the things that MPEG does not and cannot address
  - Bootstrapping licensing
  - Product interoperability
  - Compliance program
  - Education, information clearinghouse
  - Marketing, exhibitions

Membership

- 3,000 US$ per year for full membership
  - Access to meetings, documents
  - Interop Program
  - Speaking opportunities and co-exhibiting with M4IF
- 300 US$ for associate membership
  - Same rights, except voting rights

Licensing—The Responsibilities

- MPEG standardizes
  - And collects Patent Statements
  - But as a rule does not discuss licensing terms
- M4IF initiates, discusses, encourages, suggests, catalyzes, bootstraps
  - But never licenses, determines, decides, …
  - And does not recommend specific licensing terms (although it may suggest alternative approaches)
- Licensors determine, decide, sell licenses
  - In discussions with the market
  - Usually in patent pools (joint licensing schemes)

Wrap Up

Roadmap—Tuesday

	The Presidio	Pacific Heights	Haight/ Ashbury	SOMA
9:00	701 QuickTime State of the Union			
10:30			702 MPEG-4 Demystified	703 The Pro in QuickTime Pro
12:00				
2:00			704 Introduction to QuickTime Streaming	705 The Art of MPEG-4 Encoding
3:30			706 Designing a Distance Learning Solution With QuickTime	707 QuickTime Media Skins
5:00		708 QuickTime APIs Update	709 QuickTime on Campus— A Case Study	710 Preprocessing Principles

Roadmap—Wednesday

	Russian Hill	Pacific Heights	Haight/ Ashbury	SOMA
9:00	711 Audio and QuickTime		712 QuickTime for the Web	713 Introduction to QuickTime Interactivity
10:30			714 Large-Scale Webcasting: Macworld Keynote Case Study	715 QuickTime Interactivity Tools
12:00				
2:00	716 Advanced QuickTime Programming Techniques		717 3GP: Myth Versus Reality	718 AppleScript and QuickTime
3:30			719 Administering QuickTime Streaming Server	720 Content Creation for Mobile Phones: Top 10 Success Factors
5:00			721 WiFi and 3G: Let the Games Begin	722 Advanced QuickTime Interactivity

Roadmap—Thursday

	The Mission		Haight/ Ashbury	SOMA
9:00	727 QuickTime Streaming Server Programming		723 The Real Story About MPEG-4 Visual Licensing	724 What's New With QuickTime VR?
10:30			725 QuickTime and Enhanced CDs	726 Cutting-Edge Projects Using QuickTime VR
12:00				
2:00			728 Produce a Listening Party With QuickTime	729 Industrial-Strength Encoding
3:30			FF011 QuickTime Feedback Forum	
5:00				

Roadmap—Friday

	Pacific Heights	The Marina
9:00		
10:30	730 QuickTime Alternative Programming Environments	
12:00		
2:00		731 Internet Application Development: The New Frontier
3:30		
5:00		

Who to Contact

Rob Koenen
President, MPEG-4 Industry Forum
rob.koenen@m4if.org

Guillermo Ortiz
QuickTime Technology Manager
quicktimeman@apple.com

http://developer.apple.com/wwdc2003/urls.html

For More Information

For More Information

Documentation

M4IF
http://www.m4if.org

MPEG
http://mpeg.tilab.com

http://www.apple.com/mpeg4

General Information
http://www.apple.com/quicktime

Developer Information
http://developer.apple.com/documentation

QuickTime Information
http://developer.apple.com/QuickTime

Do Not Miss!

Lab: QuickTime Content Creation

Meet with QuickTime content experts for daily content specific topics that are of special interest to you. Bring your own media and PowerBook. No appointment is needed—first-come, first-served.

Sacramento Room

Lab: QuickTime Development

Meet with Apple's QuickTime engineers. Each day will emphasize a specific QuickTime development topic. Bring your own code and PowerBook. No appointment is needed—first-come, first-served.

Sacramento Room

Lab Hours

Monday	12:00pm to 6:00pm
Tuesday	9:00am to 6:00pm
Wednesday	9:00am to 6:00pm
Thursday	9:00am to 4:30pm
Friday	9:00am to 6:00pm

Q&A

Aimee Nugent
Sr. Product Manager, QuickTime
nugent@apple.com

http://developer.apple.com/wwdc2003/urls.html

WWDC 2003

Worldwide Developers Conference